وسم معجم عربي لتحليل الآراء والمشاعر

د.ابراهيم حبيبي ود.عبد الحميد الجيهاد
معهد الدراسات والأبحاث للتعريب
جامعة محمد الخامس-المغرب

الملخص:

تزامن التطور التقاني (التكنولوجي) الكبير وكذا استخدام الحواسيب على نطاق واسع في مختلف مجالات الحياة مع تطور أدوات الويب (Web 2.0)، وقد انعكست هذه الدينامية على صفحات المواقع الاجتماعية، والمنتديات وكذا مواقع البيع والشراء وغيرها من فضاءات الشبكة الإلكترونية، الشيء الذي مكن المستخدم من إبداء آرائه والتعبير عن مشاعره تُجاه قضية من القضايا المعروضة في أحد المواقع، فكانت الحاجة ماسة لتحليل هذه الآراء والمشاعر، فتبين أن تحليلها يدويا سيستغرق وقتا كبيرا، لهذا الغرض تدخل الباحثون في مجال المعالجة الآلية للغات الطبيعية للتفكير في توفير أهم الوسائل اللازمة في عملية تحليل هذه الآراء وتقييمها بشكل آلي لاستخلاص وجهات نظر المستخدم حول مختلف المواضيع المعروضة في هذه المواقع انطلاقا من الكلمات الموظفة في آرائه، لهذا الغرض تطلب توفير معجم متخصص يتم الاستعانة به في تقييم آراء ومشاعر المستخدم. من هنا برزت أهمية بناء ووسم معجم عربي باعتباره موردا وأداة أساسية في عملية تحليل الآراء والمشاعر وذلك لتجاوز قصور الأبحاث السابقة التي ركزت في عملية تحليل وجهات نظر المستعمل لهذه المواقع على ترجمة المعاجم الغربية خاصة اللغة الإنجليزية.

تعددت المقاربات المتبعة في وسم كلمات المعجم الخاص بتحليل الآراء والمشاعر، فمنها ما يتم وسمها وفق وسم ثنائي (موجب – سالب)، وآخر وفق

وسم ثلاثي (موجب – سالب – محايد)، كما أن هناك ما يوسم وفق بعض الحقول المعجمية، ومنها ما يوسم وفق ما جاءت به نظرية الأبريسال (Appraisal Theory)،حيث يُعتمد في هذه النظرية وسم الكلمات وفق حقول معجمية كبرى تنتج عنها حقول فرعية، من هذا المنطلق، سنعمل على توسيم معجم عربي تبلغ عدد دخلاته المعجمية 21453 موسومة وموزعة بين الأفعال والأسماء، ومن جهة أخرى سنضيف بعض الحقول المعجمية الفرعية إلى الحقول الرئيسة المعتمدة في نظرية الأبريسال بالاعتماد على ما جاء في الموروث العربي المعجمي المصنف وفق الحقول المعجمية قديما وحديثا، فمن القديم نجد على سبيل المثال مُؤلف: "نجعة الرائد وشرعة الوارد في المترادف والمتوارد" لإبراهيم اليازجي اللبناني، أما الحديث منها نجد على سبيل المثل معجم "المكنز الكبير" لأحمد مختار عمر، ثم نسم هذه الكلمات حسب حقولها المعجمية مع تحديد الموجب والسالب والمحايد من هذه الكلمات كما سنحدد درجة قوتها ما إذا كانت مرتفعة، متوسطة، أو منخفضة بالاعتماد على بعض المعاجم العربية، ليتم وضعه مفتوح المصدر فيوظف في البرامج التي تعمل على تحليل الآراء والمشاعر باللغة العربية.

تقديم:

تعتبر اللسانيات الحاسوبية أحد فروع علم اللسانيات التطبيقية، إذ تعد علما تطبيقيا حديثا يقوم على استغلال التقانة (التكنولوجيا) المتطورة التي لها صلة ببرمجة الأنظمة المعلوماتية الخاصة بالمعالجة الآلية لمختلف اللغات الطبيعية، الشيء الذي ساهم في نمذجة مختلف المستويات اللسانية ومحاولة برمجتها، وقد تعددت مجالات تطبيقات اللسانيات الحاسوبية، وتم تتويج ذلك بمرحلة صياغة تقانية (تكنولوجية) خاصة تمثلت في تكنولوجية اللغة، وقد اتخذت مجالات تطبيق اللسانيات والتقنيات الحاسوبية العديد من البرامج والأنظمة على سبيل المثال لا الحصر: النظم الآلية للتدقيق الهجائي والنحوي، الاستخلاص الآلي، الفهرسة الآلية، برامج تعلم اللغات وتعليمها، الترجمة الآلية، نظم استرجاع المعلومات والمعاجم والقواميس الإلكترونية[1].

لقد أدى تطور أدوات الويب (Web 2.0)، إلى المساهمة في إحداث مواقع الشبكة العنكبوتية لغرض تحقيق العملية التفاعلية بين المستخدم والمعلومات المنشورة في المواقع، الشيء الذي ساهم في تطوير المواقع الاجتماعية، والمنتديات ومواقع البيع والشراء، وغيرها، فأصبح المستخدم بإمكانه قراءة المعلومات المنشورة وإبداء آرائه والتعبير عن مشاعره تُجاهها، الأمر الذي سيطرح إشكال تحليل هذه الآراء وتقييمها بهدف معرفة رغبات المستخدم، فلوحظ أن تحليلها يدويا سيستغرق وقتا طويلا، لهذا الغرض تدخل الباحثون في مجال المعالجة الآلية للغات الطبيعية، قصد توفير أهم الوسائل اللازمة في عملية تحليل هذه الآراء وتقييمها بشكل آلي لاستخلاص وجهات نظر المستخدم حول مختلف المواضيع المعروضة في هذه المواقع انطلاقا من الكلمات الموظفة في آرائه، لهذا الغرض تطلب توفير معجم متخصص يتم الاستعانة به في تقييم آراء ومشاعر المستخدم، من هنا برزت أهمية بناء ووسم معجم عربي إذ يعتبر موردا وأداة أساسية في عملية تحليل الآراء والمشاعر وذلك لتجاوز قصور الأبحاث العربية السابقة

(1) س. باية، "اللسانيات الحاسوبية والمعجمية العربية"، *لغة – كلام*، ص 228–241، 2017م-1438هـ.

المرتكزة على ترجمة المعاجم الغربية، فقد بُني عملنا على مجموعة من الدوافع الوجيهة حيث تُرد إجمالا إلى:

- ✓ **قلة المعاجم الخاصة بمعالجة الآراء والمشاعر باللغة العربية ومحدوديتها بالمقارنة مع الموروث المعجمي العربي.**
- ✓ **أغلب المعاجم المتوفرة المتعلقة بتصنيف الآراء والمشاعر المكتوبة باللغة العربية تم بناؤها اعتمادا على الترجمة.**
- ✓ **لا تشتمل المعاجم المترجمة إلى اللغة العربية على أقسام الكلام العربي والتي توظف في التعبير عن الآراء والمشاعر.**
- ✓ **أغلب المعاجم المتوفرة حاليا تعتمد على تصنيفات ضيقة: تصنيف ثنائي (موجب – سالب) أو ثلاثي (موجب – سالب – محايد) على الأكثر.**

نهدف من خلال هذا العمل إذن إلى تقديم الخطوات المتبعة في بناء معجم عربي متخصص، يضم عددا كبيرا من أقسام الكلام العربي (الأسماء والأفعال والحروف) التي يمكن أن توظف في عملية معالجة الآراء والمشاعر معالجة آلية وتصنيفها وفق الحقول المعجمية المتبناة من طرف الباحثين "مارتن وويت" (Martin and White) على ضوء نظرية الأبريسال[2] من جهة وما تمت إضافته من حقول معجمة جديدة من جهة أخرى.

لتحقيق هذا الهدف سنقوم بتقديم تعريف لمجال تحليل الآراء والمشاعر مع الإشارة إلى أهم العناصر المشكلة لعملية المعالجة الآلية للآراء والمشاعر والإشارة إلى مختلف المستويات التي تشملها هذه العملية، ثم سنقف بشكل مقتضب على أهم المقاربات المعتمدة في تحليل الآراء والمشاعر، لنعرج على الحديث عن أنواع معاجم تحليل الآراء والمشاعر باعتبار اللغات التي تشتمل عليها، ثم الإشارة إلى أهم المقاربات المتبعة في بناء وتصنيف هذا النوع من المعاجم، بعد ذلك سنسلط الضوء على أشهر المعاجم المتخصصة في مجال تحليل

(2) J. R. Martin و P. R. R. White, *The language of evaluation: appraisal in English*. New York: Palgrave Macmillan, 2005

الآراء والمشاعر خاصة المترجم منها إلى اللغة العربية، ثم نقدم تعريفا بنظرية الأبريسال مع الإشارة إلى أهم النظم التي تنبني عليها والكشف عن المقاربة المعتمدة من طرف "مارتن وويت" في تصنيف المعجم، لنفصل الحديث بعد ذلك في الخطوات المتبعة في بناء معجم عربي موجه لتحليل الآراء والمشاعر الخاص باللغة العربية، لنختم عملنا المتواضع باستنتاج خلاصات وكذا رهانات البحث المستقبلية.

1. التعريف بمجال تحليل الآراء والمشاعر:

يعتبر تحليل المشاعر، كما سبقت الإشارة إلى ذلك، مجال بحث يعمل على معالجة الآراء، المشاعر، التقييمات، التثمينات، مواقف وعواطف الإنسان حول بعض الأشياء سواء كانت اتجاهات أو منتوجات، أو خدمات، أو تنظيمات، أو أحداث، أو مواضيع معينة...، ومن خلال هذا التعريف فإن عملية معالجة الآراء والمشاعر تهم خمسة عناصر على الشكل الآتي:

– الموضوع (Object): هو الشيء الذي أُعطي فيه الرأي سواء كان منتوجا أو خدمة أو قضية أو موضوعا أو شخصية أو تنظيما أو غيرها.

– الهيئة (Aspect): هي المسؤولة عن الموضوع الذي أُعطي فيه الرأي المعني بالرأي، المقصود هنا الجهة المعنية بالتعاليق التي نشرت حولها.

– موجهات المشاعر (Sentiment Orientation): المؤشر الذي يحيل على أن الرأي إما موجبا، سالبا، أو محايدا.

– صاحب الرأي (Opinion Holder): يهم ما قدم الرأي حول الموضوع سواء تعلق الأمر بشخص أو هيئة.

– الوقت (Time): الوقت الذي نُشر فيه التعليق حول الموضوع[3].

ويمكن تحليل الآراء والمشاعر وفق ثلاثة مستويات تتمثل في:

(3) N. Boudad, R. Faizi, R. Oulad Haj Thami, و R. Chiheb, "Sentiment analysis in Arabic: A review of the literature", *Ain Shams Eng. J.*,.2018 م 9، عدد 4، ص 2479–2490، ديسمبر

- **التحليل على مستوى الوثيقة** (Document Level): يهتم هذا المستوى بتصنيف الوثيقة حسب ما تضمنته من آراء إلى وثيقة موجبة أو سالبة أو محايدة، إذ يعتبر هذا المستوى الوثيقة أنها تمثل رأيا واحدا، مما يصعب تطبيقه عن الوثائق التي تتحدث عن مقارنة شيء بشيء آخر.

- **التحليل على مستوى الجملة**: (Sentence Level) يعمل هذا المستوى على تصنيف الجمل - حسب الرأي الذي عبرت عنه - إلى إيجابي، سلبي، أو محايد، ويهتم هذا المستوى بالتصنيف الذاتي الذي تعبر فيه الجملة عن وجهات نظر وآراء ذاتية، كما يهتم بالتصنيف الموضوعي الذي تعبر فيه الجملة عن معلومات واقعية، وهنا ينبغي الإشارة إلى أن الجمل الذاتية لا تعادل الجمل المعبرة عن المشاعر، كما أن الجمل الموضوعية من شأنها أن تحيل على رأي ولو بشكل ضمني مثل: "اشترينا السيارة الشهر الماضي وقد سقطت لها ممسحة الزجاج الأمامي؛ فهذا المثال يحيل على رأي يقضي الحكم برداءة نوع السيارة وإن كان ضمنيا، ويعد هذا المستوى من التحليل بمثابة المحدد الأساسي لتوجيه الكلمة باعتماد السياق اعتمادا أساسيا، ويشمل التصنيف حسب هذا المستوى جمل المقارنة وأسلوب المدح والذم إلى غيرها من الأساليب. ورغم ذلك فقد أشار الباحثون أن مستوى الجملة ما زال يحتاج إلى المزيد من الدراسات.

- **التحليل على مستوى الكلمة** (Entity and Aspect level): يعتبر هذا المستوى أكثر تدقيقا في عملية التحليل؛ إذ إن الغرض منه هو الكشف عن العناصر الخمسة السابقة الذكر أو البعض منها، وذلك عبر مهمتين أساسيتين هما: استخراج الكلمة أو العبارة الدالة على الآراء والمشاعر ثم تصنيفها حسب توجيهها، ما إذا كانت إيجابية أو سلبية أو محايدة[4].

2. المقاربات المعتمدة في معالجة الآراء والمشاعر:

تتمثل المقاربات المعتمدة في مجال معالجة الآراء والمشاعر في المقاربة اللسانية ومقاربة تعليم الآلة والمقاربة الهجينة، وفيما يلي سنقدم ملخصا لمقاربة

(4) B. Liu, "Sentiment Analysis and Opinion Mining",.168 ص

تعليم الآلة والمقاربة الهجينة بشكل عام، وسنتحدث بشكل خاص عن المقاربة اللسانية.

أ. مقاربة تعليم الآلة (Machine Learning Approach):

هذه المقاربة هي أكثر المقاربات استعمالا في تصنيف الآراء والمشاعر، وتشتمل على:

- **التعليم الموجه** (Supervised Learning).
- **التعليم نصف الموجه** (semi-supervised learning).
- **التعليم غير الموجه** (Unsupervised Learning).

التعليم الموجه: تعتمد مقاربة التعليم الموجه على معطيات موسومة يدويا ليتم بعد ذلك وسم معطيات أخرى آليا، وكمثال على ذلك، إذا كان عندنا تصنيف مجالات في مكتبة معينة، فإنه يتم على إثر هذا التصنيف، تصنيف مستندات جديدة حسب التخصصات الموجودة في التصنيفات المسبقة، ومن بين اللوغاريتمات المعتمدة في هذا التعلم نجد:

- Trees of Decisions ;
- Naïve Bayes (or Simple Bayes) ;
- Neural Networks ;
- Vector Support Machines (or SVM) ;
- Genetic Programming.

وقد تم اعتماد هذه المقاربة من طرف العديد من الباحثين من بينهم "بينغ وفيتينغ (S. Vaithyanathan et Pang, L. Lee)[5]"ووإنسون وفليت (E. Eensoo et M. Valette)[6].

(5) B. Pang, L. Lee,و S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques", 2002,.86–79 ص ,10 م.

(6) E. E. M. Valette, "Sur l'application de méthodes textométriques à la construction de critères de classi!cation en analyse des sentiments", 8 ص.

التعليم نصف الموجه: تعمل مقاربة تعليم نصف الموجه على الجمع بين التعليم الموجه وغير الموجه، أي الاعتماد على المعطيات الموسومة وغير الموسومة لتحسين نوعية التصنيف ويمكن الإحالة على بعض الأبحاث التي اعتمدت هذه المقاربة مثل "عبد المجيد وديـاب (M. Abdul-Mageed et M. T. Diab)[7] "وارتغوسا ومرتن وكرلوس (A. Ortigosa, J. M. Martín, et R. M. Carro)"[8].

التعليم غير الموجه: عكس المقاربتين السابقتين؛ فمقاربة التعليم غير الموجه لا تعتمد في تصنيفها على معطيات موجودة مسبقا، ومن بين اللوغاريتمات المعتمدة في هذا المقاربة نجد:

- Clustering Using Representatives Algorithm (CURE).
- Balanced Iterative Reducing and Clustering Using Hierarchies Algorithms (BIRCH).
- Robust Clustering Using Links Algorithm De (ROCK).
- Tree Structured Vector Quantization Algorithm (TSVQ).

ويمكن الإحالة على بعض الأعمال التي اعتمدت هذه المقاربة، مثل "بردي والحداد (Brody et N. Elhadad)"[9] ، "وفريمندس وآخرون (Fernández and all)"[10].

ب. المقاربة اللسانية (Linguistic Approach):

تعددت الأسماء في هذه المقاربة حسب ما يوظف من المستويات اللسانية، بل هناك من يسميها بالمقاربة الرمزية أو المقاربة المعجمية وكلها مفاهيم تحيل على الاعتماد على المستويات اللسانية في معالجة الآراء والمشاعر، لكن الأهم فيها هو:

(7) M. Abdul-Mageed و M. Diab, "AWATIF: A Multi-Genre Corpus for Modern Standard Arabic Subjectivity and Sentiment Analysis",.8 ص

(8) A. Ortigosa, J. M. Martín, و R. M. Carro, "Sentiment analysis in Facebook and its application to e-learning", *Comput. Hum. Behav.*,.2014 فبراير ,541–527 ص ,31 م

(9) S. Brody و N. Elhadad, "An Unsupervised Aspect-Sentiment Model for Online Reviews", ص 9.

(10) M. Fernández-Gavilanes, T. Álvarez-López, J. Juncal-Martínez, E. Costa-Montenegro, و F. Javier González-Castaño, "Unsupervised method for sentiment analysis in online texts", *Expert Syst. Appl.*,.2016 أكتوبر ,75–57 ص 58 م

أولا: المستوى المعجمي المتمثل في توفير دخلات معجمية مصنفة إلى موجبة وسالبة ومحايدة بطريقة يدوية أو نصف آلية، وسنوضح هذه المقاربات المعتمدة في بناء معجم خاص بالتعبير عن الآراء والمشاعر في الفقرة المتعلقة بالمقاربات المعتمدة في بناء وتصنيف المعاجم.

ثانيا: المستوى التركيبي المتمثل في الأنظمة التي تعمل على تقطيع النصوص إلى وحدات بعدها يتم الانتقال إلى المعجم المعتمد لتقديم نتيجة التحليل حسب الكلمات الموظفة في التعاليق.

كما أن بعض الأبحاث تعتمد كذلك إلى جانب المستويين السابقين على المستوى النحوي المتمثل في وضع مجموعة من القواعد النحوية مثل النفي والنهي وغيرها من القواعد التي يمكن اعتمادها من طرف المستعمل في التعبير عن الآراء والمشاعر، كما يلعب المستوى الصرفي دورا كبيرا في التعرف على السوابق واللواحق والأحشاء المرتبطة بالكلمات المكونة للنص أو التعليق. وفي هذا الصدد يمكن الإحالة إلى بعض الأعمال التي اعتمدت هذه المقاربة لتحليل الآراء والمشاعر مثل: "البحيري والحمود واللويد (Alhumoud, T. Albuhairi, et W. Alohaideb)"[11]

كما أن هناك من يشتغل وفق نظرية الأبريسال، وهي نظرية لسانية وظيفية اعتمدها كل من "مارتن وويت" لتقييم نصوص اللغة الإنجليزية بالاعتماد على ثلاثة أنظمة، وهي: **الموقف والارتباط والتدرج**، لكن الأبحاث الحاسوبية اهتمت فقط بالنظام الأول باعتباره النظام الذي يشكل المعجم المعتمد في مجال تحليل الآراء والمشاعر، حيث صنف الباحثان مارتن وويت نظام الموقف إلى ثلاثة حقول معجمية كبرى، وهي: **الانفعال، والحكم، والتقدير**، ويضم كل حقل رئيسي حقولا فرعية أخرى. ويساعد الاعتماد على هذا المعجم وفق الحقول

(11) S. Alhumoud, T. Albuhairi, و W. Alohaideb, "Hybrid sentiment analyser for Arabic tweets using R", في *7 2015 th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K)*, 2015, م 01, ص 417–424.

المعجمية التي يتضمنها على تقديم أدق لما اشتمل عليه التعليق أو النص من آراء ومشاعر[2]. وقد اعتمد العديد من الباحثين هذه المقاربة في معالجتهم للآراء والمشاعر الواردة في تعاليق مأخوذة من مواقع تفاعلية متعددة المجالات، ومن بينهم "فيراري وأخرون (Ferrari and all)"[12] "وكرول وريد (Read et J. Carroll)"[13] "وكردين (P. Gardin)"[14] "وبلوم (K. Bloom)"[15]" "وويتلاو وآخرون (C. Whitelaw)"[16].

ت. المقاربة الهجينة (Hybrid Approach):

وهي المقاربة التي تجمع بين المقاربة اللسانية ومقاربة تعليم الآلة؛ حيث تعمل على الجمع؛ إما بين اللسانيات باعتبارها مكملة لتعليم الآلة، أو تعليم الآلة باعتباره مكملا للسانيات أو التوفيق بينهما؛ فالطريقة الأولى تتضمن استعمال الأدوات اللسانية لإعداد المدونة قبل تصنيف النصوص بمساعدة أدوات تعليم الآلة، إذ تمكن هذه المقاربة من الاحتفاظ بقوة التعليم الآلي، وتوجيه أساس التدريبات في الترتيب اليدوي المعتمد في المقاربة اللسانية في الوقت نفسه[17].

3. أنواع المعاجم:

من خلال الاطلاع على مجموعة من الأبحاث المهتمة ببناء المعاجم الموجهة لتحليل الآراء والمشاعر تبين أنه يمكن تصنيف المعاجم المتخصصة بهذا

(12) S. Ferrari, Y. Mathet, T. Charnois, و D. Legallois, "Analyse d'opinion : discours évaluatif et classification de documents", *Actes L'atelier FODOP*, م 8, ص 23–36.

(13) J. Read و J. Carroll, "Annotating expressions of Appraisal in English", *Lang. Resour. Eval.*, م 46, عدد 3, ص 421–447, سبتمبر 2012.

(14) P. Gardin, "Application de la théorie de l'Appraisal à l'analyse d'opinion", ص 4.

(15) K. Bloom, *Sentiment analysis based on appraisal theory and functional local grammars*. Illinois Institute of Technology Chicago, US, 2011.

(16) C. Whitelaw, N. Garg, و S. Argamon, "Using appraisal groups for sentiment analysis", في *Proceedings of the 14th ACM international conference on Information and knowledge management*, 2005, ص 625–631.

(17) S. Maurel, "Classification d'opinions par méthodes symbolique, statistique et hybride", 2007. , ص 7

المجال باعتبار اللغة التي تشتمل عليها إلى أربعة أنواع كبرى من المعاجم، وهي: معجم أحادي اللغة، معجم ثنائي اللغة، معجم متعدد اللغات، ومعجم متعدد المجالات، وفيما يلي تفصيل لكل نوع من هذه المعاجم:

– **المعاجم الأحادية اللغة**: يرتكز المعجم الأحادي اللغة على ضم مجموعة من الكلمات أو العبارات الدالة على الآراء والمشاعر للغة معينة، وقد شهدت اللغات الهندوأوروبية تقدما كبيرا في هذا المجال، بل تعتبر السباقة لإنشاء معاجم لتحليل الآراء والمشاعر أحادية اللغة، خاصة اللغة الإنجليزية منها، والتي تندرج ضمنها المعاجم التي تضم كلمات خاصة بلهجة بعض المناطق، وهذا ما ارتكزت عليه بعض الدراسات العربية[18].

– **المعاجم الثنائية اللغة**: اهتمت المعاجم ثنائية اللغة التي لها علاقة بمجال تحليل الآراء والمشاعر على إنشاء معجم يضم كلمات أو تعابير تستعمل في التعبير عن المواقف والمشاعر والآراء، يبدو ذلك جليا في الأعمال العربية، إذ تجد كثيرا منها ركزت على بناء معاجم تشتمل على كلمات وتعابير اللغة العربية الفصحى وأخرى عامية تهم منطقة أو دولة معينة، كما توجد بعض الأعمال التي ارتكزت على بناء معجم ثنائي يضم اللغة العربية واللغة الإنجليزية[19].

– **المعاجم المتعددة اللغات**: نظرا لكون المعلقين في كثير من الأحيان يستعملون أكثر من لغة، كانت الحاجة ماسة لإنشاء معجم متعدد اللغات، يضم أكثر من لغتين، وذلك لتحليل كل ما اشتملت عليه الآراء والتعابير عن مشاعر بلغات مختلفة، وترتكز عملية توفير هذا النوع من المعاجم أساسا على الترجمة الآلية لمعجم من المعاجم[20].

---

(18) F. H. H. Mahyoub, M. A. Siddiqui, و M. Y. Dahab, "Building an Arabic Sentiment Lexicon Using Semi-supervised Learning", *J. King Saud Univ. - Comput. Inf. Sci.*, م 26, عدد 4, ص 417–424, ديسمبر 2014.

(19) S. R. El-Beltagy و A. Ali, "Open issues in the sentiment analysis of Arabic social media: A case study", *9 في 2013 th International Conference on Innovations in Information Technology (IIT)*, 2013, ص 215–220.

(20) S. L. Lo, E. Cambria, R. Chiong, و D. Cornforth, "Multilingual sentiment analysis: from formal to informal and scarce resource languages", *Artif. Intell. Rev.*, م 48، عدد 4، ص 499–527، ديسمبر 2017

– **المعاجم المتعددة المجالات**: جاءت فكرة وضع معجم متعدد المجالات انطلاقا من بناء مدونة متعددة المجالات، إذ تعتبر في كثير من الأحيان كلمة موجبة في مجال معين ولكنها سالبة في مجال آخر؛ فإذا أخذنا كلمة "طويلة" على سبيل المثال فتوجيهها يختلف من مجال إلى مجال، ففي جملة "مسافة السفر طويلة" يعتبر توجيه الكلمة سالبا، في حين إذا قلنا "تشتغل هذه البطارية مدة طويلة" فتوجيه هذه الكلمة يعتبر موجبا[21].

4. المقاربات المعتمدة في بناء معجم خاص بتحليل الآراء والمشاعر:

يعتبر بناء المعاجم المتخصصة في مجال تحليل الآراء والمشاعر من بين الاهتمامات البحثية في مجال المعالجة الآلية للغات الطبيعية، وقد تعددت المقاربات المعتمدة في بناء مثل هذه المعاجم وتصنيفها، ونقدم بإيجاز في هذا الصدد المقاربات المعتمدة في بناء هذا النوع من المعاجم على الشكل الآتي:

أ. المقاربة اليدوية (manual approach):

تعتبر من بين المقاربة البسيطة والناجعة لبناء المعاجم، حيث تنبني على المعرفة الإنسانية المتمثلة في جمع الكلمات التي لها دلالة على الآراء والمشاعر ثم تحديد قطبية كل كلمة، ورغم نجاعته، إلا أنه منهج مكلف، حيث يتطلب وقتا كبيرا لحصر كل الكلمات التي لها علاقة بتحليل الآراء والمشاعر، لهذا تم التفكير في استبداله بمناهج آلية وشبه آلية.

ب. مقاربة المدونة (corpus approach):

تنطلق هذه المقاربة من مجموعة من الكلمات المنفصلة (germs) التي لها قطبية معروفة، ويتم حساب مماثلة الكلمات باستعمال التحليل الإحصائي لسياقات ترابط الكلمات فيما بينها لسانيا أو دلاليا بالاعتماد على نماذج متجهات المسافات (Vector Spaces) التي تقدم المحتوى الدلالي للكلمات على شكل

(21) H. ElSaharو S. R. El-Beltagy, "Building Large Arabic Multi-domain Resources for Sentiment Analysis", في *Computational Linguistics and Intelligent Text Processing*, 2015,.34–23 ص

متجهات، فيتم اعتبار الكلمة موجبة إذا كانت قريبة من الكلمة الأصل، وتكون سالبة إذا كانت بعيدة عنها، وقد طبقت هذه المقاربة على مجموعة من المدونات الكبيرة الحجم.

ت. مقاربة تعتمد على القاموس (Dictionary-Based Approaches):

ث. يُعتمد في هذه المقاربة على قواميس أو موارد لسانية للمرادفات والاضداد مثل "WordNet" لتحديد التوجيه الدلالي لكلمات جديدة، فانطلاقا من كلمات منفصلة لها قطبيتها، يمكن الحصول على كلمات جديدة لها نفس قطبية الكلمات الأصلية باستعمال لوغاريتم بوتسترابين (bootstrapping) التي تمكن من ربط الاتصال بالمرادفات والاضداد للمورد اللساني.

أما فيما يخص تصنيف كلمات معجم تحليل الآراء والمشاعر فقد تعددت طرق تصنيفها، وسنوضحها مختصرة في النقاط الآتية:

- تصفيف المعجم وفق قطبية ثنائية: حيث يتم تصنيف كلمات المعجم إلى كلمات موجبة وكلمات سالبة.
- تصفيف المعجم وفق قطبية ثلاثية: حيث يتم تصنيف كلمات المعجم إلى كلمات موجبة وكلمات سالبة وكلمات محايدة.
- تصنيف المعجم وفق درجة القطبية: صنفت فيه الكلمات إلى موجب قوي، وموجب ضعيف وسالب قوي وسالب ضعيف.
- تصنيف وفق الحقول المعجمية: ويتم في هذا النوع من التصنيفات اعتماد مجموعة من الحقول المعجمية لها ارتباط بالآراء والمشاعر.

وقد اخترنا اعتماد التصنيف وفق الحقول المعجمية في سيرورة بناء معجمنا، وذلك راجع إلى الدقة التي يتميز بها عن التصنيفات الأخرى، ولأنه يساعد على تحديد وجهة نظر مستعمل المواقع تحديدا دقيقا.

5. أهم المعاجم المتوفرة:

يعتبر المعجم من أهم الموارد التي ينبغي توفيرها للخوض في مجال التحليل الآلي للآراء والمشاعر، وهذا ما نجده ضمن مجموعة من الدراسات التي تهتم ببناء المعاجم التي يتم توظيفها في عملية التحليل، والباحث في هذا المجال سيجد أن اللغات الأخرى قد حققت تقدما بارزا في توفير هذا النوع من المعاجم، بل تجاوزوا ذلك إلى إعداد البرامج المعتمدة في تحليل الآراء والمشاعر وفق مقاربات مختلفة، أما بالنسبة للغة العربية، فالمتتبع للدراسات المنجزة في هذا المجال، سيجد أن أغلب المعاجم المعتمدة مترجمة من اللغات الأجنبية خاصة اللغة الإنجليزية باعتبارها اللغة التي حققت تقدما بارزا في هذا المجال، كما نجد بعض الباحثين وضعوا معاجم تشتمل على كلمات اللغات المحلية، وسنحاول في هذا المحور الإشارة إلى أغلب المعاجم المتوفرة في هذا المجال والإشارة إلى المترجم منها إلى اللغة العربية، ونشير كذلك إلى المعاجم التي تشتمل على اللغة المحلية.

ج. معجم المجلس الوطني للأبحاث في كندا " Conseil national de recherches Canada**"** (NRC):

تم إنشاء هذا المعجم من طرف هيئة بالمجلس الوطني للبحوث بكندا وهو معجم يضم كلمات إنجليزية بلغ عددها 14182 دخلة معجمية ارتكزت على الصفات، تم تصنيفها يدويا وفق ثمانية حقول معجمية مرتبطة بالمشاعر، هي: **"الغضب والترقب، الاشمئزاز والخوف والفرح والحزن والدهشة والثقة"** مع تحديد الموجب منها والسالب.

إن ما يميز هذا المعجم كونه ترجم إلى العديد من اللغات ناهزت أربعين لغة بالاعتماد على المترجم الآلي (Google traduction)، من بينها اللغة العربية التي أشرف على ترجمتها الباحث محمد سيف ويمكن تحميله من هذا الرابط، ومن اللغات التي ترجم إليها هذا المعجم: "الكتالانية، الصينية، الدانماركية، الإسبانية، الفنلندية، الفرنسية، الألمانية، اليونانية، العبرية، الهندية، الإيرلندية، الإيطالية،

اليابانية، اللاتينية، الفارسية، البرتغالية، الرومانية، الروسية، الصومالية، السودانية، السويدية، التايلاندية، التركية، والأوكرانية".

أ. المعجم الفرنسي الموسع للمشاعر "A French Expanded Emotion Lexicon" (FEEL)

بني هذا المعجم أساسا بالاعتماد على ترجمة معجم المجلس الوطني للبحوث بكندا وذلك بتوظيف ستة مواقع للترجمة، وهي: " Google Translate, Bing Translate, Collins Translator, Reverso, Bab.la et Word Reference"، ثم قام الفريق بإغناء المعجم بالبحث عن مرادفات كل كلمات المعجم في ثماني مواقع هي: Reverso Dictionary, Bab.la, Atlas, Thesaurus, Ortolang, SensAgent , The Free Dictionary, and the Synonym website، فبعد حذف الكلمات المتماثلة، تم الوصول إلى 141428 دخلة معجمية، تمثلت في الصفات، ولجعل هذا المعجم ذا قيمة علمية تم التعاقد مع أحد الأشخاص المتخصصين في مجال الترجمة، وبعد مدة دامت شهرين من المراجعة تم تقليص المعجم إلى 81757 دخلة معجمية، لم يوسم منها إلا 14129[22] دخلة معجمية مقسمة بين 11979 كلمة مفردة و2148 كلمة مركبة[23] ، وتجدر الإشارة إلى أن هذا المعجم متوفر للتحميل وهو على صيغة "Excel" ويمكن تحميله من الرابط الآتي[24].

ب. معجم منظور متعدد سؤال جواب " Multi-Perspective Question Answering" (MPQA):

تم إنشاء هذا المعجم من طرف كل من: تيريزا ويلسن، جنسي ويب وبوول هُوفْمن[25]، MPQA هي اختصار لمنظور متعدد سؤال جواب وهو معجم

(22) "NRC Emotion Lexicon". موجود في: [مباشر على الإنترنت]. https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm.
[تاريخ الوصول: 21-يونيو-2019].

(23) A. Abdaoui "وآخرون ADVANSE: Sentiment, Opinion and Emotion Analysis in French Tweets", في *DEFT: Défi Fouille de Texte*, Caen, France, 2015.

(24) http://www.lirmm.fr/~abdaoui/FEEL.csv

(25) A. Abdaoui, J. Azé, S. Bringay, و P. Poncelet, "FEEL: a French Expanded Emotion Lexicon", *Lang. Resour. Eval.*, سبتمبر 2017. م 51، عدد 3، ص 833–855.

إنجليزي الأصل يضم 8214[26] كلمة موزعة بين الأسماء والصفات وبعض الأفعال صنفت وفق التصنيف الثلاثي [موجب، سالب، محايد]، وقد تُرجمت نسبة منه إلى اللغة العربية من طرف بعض الباحثين ثم قام فريق المعالجة الآلية للغات الطبيعية بجامعة صفاقس بتونس بترجمة الكلمات المتبقية، وإغنائه بألفاظ أخرى واعتمدوا في ترجمة هذا المعجم على ArabicWordNet و[27]Sentistrength، لكنهم أغفلوا توسيم الكلمات كما كانت محددة في النسخة الإنجليزية، بالإضافة إلى ذلك فقد تم تصنيف الكلمات حسب ما يلي: "موجب قوي – موجب ضعيف – سالب قوي – سالب ضعيف".

ت. معجم (SentiWordNet):

يعد هذا المعجم من بين أشهر المعاجم المستعملة في مجال تحليل الآراء والمشاعر، صنفت كلماته وفق التصنيف الثنائي [موجب–سالب]. ولهذا المعجم أربعة إصدارات، هي:

الإصدار الأول: SENTIWORDNET 1.0 قُدم من طرف الباحثين إيزيلي وسبستياني (Esuliet Sebastiani) كورقة بحث أولية قدمت في مؤتمر الموارد المعجمية لتحليل الآراء وجعله رهن إشارة الأبحاث المتعلقة بمجال تحليل الآراء والمشاعر سنة 2006.

- **الإصدار الثاني**: SENTIWORDNET 1.1 قدم من طرف الباحثين إيزيلي وسبستياني في إطار تقرير تقني في معهد علوم وتكنولوجيا المعلومات، التابع للمجلس الوطني للبحوث، بيزا سنة 2007[28] .

- **الإصدار الثالث**: SENTIWORDNET 2.0 عرضه الباحث إيزيلي في إطار أطروحته بعنوان: "توليد الآلي للموارد المعجمية الخاصة بتحليل الآراء"

(26) http://mpqa.cs.pitt.edu/

(27) A. Abdaoui, J. Azé, S. Bringay,و P. Poncelet, "FEEL: a French Expanded Emotion Lexicon", *Lang. Resour. Eval.*,.2017 سبتمبر ,855–833 ص ,3 عدد ,51 م

(28) T. Wilson, J. Wiebe,و P. Hoffmann, "Recognizing Contextual Polarity in Phrase-level Sentiment Analysis", في *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA, 2005,.354–347 ص

(Automatic Generation of Lexical Resources for Opinion Mining) أطروحته سنة 2008.

- **الإصدار الرابع**: SENTIWORDNET 3.0 قدم من طرف إيزيلي وآخرين سنة 2010[29].

ترجم هذا المورد المعجمي إلى العديد من اللغات من بينها اللغة العربية من طرف الباحث سماح الحازم وآخرين سنة 2013 سمي "ArabicSentiWordNet" وهو عبارة عن إصدار من بين الإصدارات الناتجة عن "EnglishSentiWordNet"، وذلك باتباع ثلاثة خطوات رئيسة:

1. **تحديث إصدار** arabeWordNet 2.0 **للحصول على الإصدار الثالث**: arabeWordNet 3.0، عمل الباحثون اعتمادا على مقارنة، على إغناء الإصدار الثاني للغة العربية بإضافة ما جد في الإصدار الثالث للغة الإنجليزية (EnglaisWordNet 3.0).

2. **البحث عن النسخة العربية** (ArabeSentiWordNet3.0) **في النسخة الإنجليزية** (EnglishSentiWordNet0.3).

3. **مراجعة النسخة العربية** (ArabeSentiWordNet3.0): همت هذه العملية بالأساس الاحتفاظ بالحقول المتعلقة بالمشاعر الموجودة في النسخة الإنجليزية وحذف الزائد من النسخة العربية[30].

ث. معجم مرتن وويت:

ترجم هذا المعجم إلى اللغة العربية من طرف بعض الباحثين العرب، منهم الباحثة الجزائرية غيتة يمينة تليلي التي قامت بترجمة 325 كلمة مأخوذة من كتاب الباحثين مرتن وويت، وتم تطبيقها في معالجة الآراء في مجال الاقتصاد[31].

(29) A. Bayoudhi, H. Koubaa, H. Ghorbel, و L. Hadrich Belguith, "Vers un lexique arabe pour l'analyse des opinions et des sentiments", في *Proc of the 5th International Conference on Arabic Language Processing CITALA*, 2014, م 14.

(30) A. Esuli و F. Sebastiani, "SENTIWORDNET: A high-coverage lexical resource for opinion mining", *Evaluation*, يناير 2007.

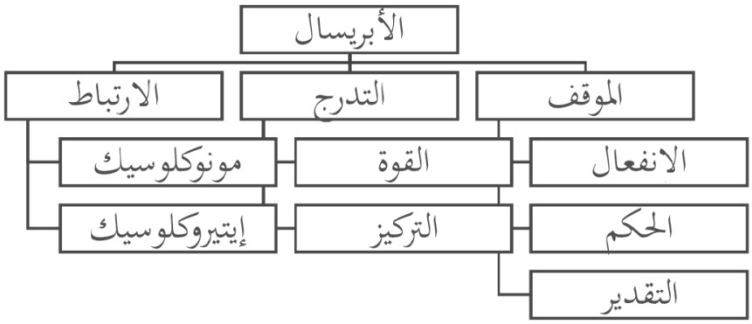
(31) A. Esuli, "Automatic Generation of Lexical Resources for Opinion Mining: Models, Algorithms and Applications", *ACM SIGIR Forum*, م 42, عدد 2, ص 2, 2008.

في حين قام الباحث المغربي أحمد عمار بنفس العمل مع إغنائه بالاعتماد على (WordNet Arabic وBabelNet) ليصل عدد دخلات المعجم 1400 كلمة[32]، وهذا قليل بالمقارنة مع موروث اللغة العربية.

6. نظرية الأبريسال

تعد نظرية الأبريسال من نظريات علم النفس التي تهتم بالعواطف المستنتجة والتي يتم تقييمها أو تقديرها انطلاقا من ردود أفعال البشر المختلفة. وينتج هذا التقييم أساسا بسبب استجابة عاطفية أو وجدانية، فهي نظرية تهتم بتعبيرات المتكلم عن العواطف والآراء ومشاعره، وقد ارتبطت– بفضل جهود الباحثين "مارتن وويت" في إطار إحيائها- ارتباطا وثيقا في الأبحاث الأخيرة بنظام اللسانيات الوظيفية خاصة مع هاليداي[33]. وقدتمثل ذلك في الاهتمام بمستويين لسانيين، هما: المستوى المعجمي والمستوى التركيبي، لغرض دراسة النصوص الإنجليزية المشتملة على الآراء والمشاعر.

وجه الباحثان مرتن وويت مسار نظرية الأبريسال إلى تحليل الآراء والمشاعر في النصوص الإنجليزية بالاعتماد على ثلاثة أنظمة أساسية، تتمثل في: "الموقف – الارتباط – التدرج"، تتفرع عن هذه الأنظمة العديد من المكونات التي تساعد على فهم الآراء والمشاعر وتحليلها بطريقة دقيقة[2] (انظر الخطاطة 1)



(32) S. Alhazmi, "Arabic SentiWordNet in Relation to SentiWordNet 3.0",.2013 ,11 ص

(33) G. Y. Tlili, "L'application de la théorie de l'Appraisal pour l'identification de l'opinion dans le texte arabe."

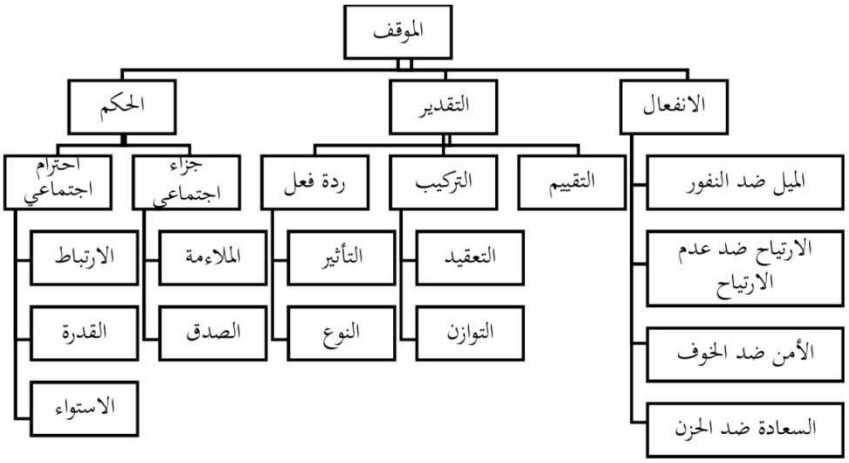
- **الموقف**: هو نسق معجمي رئيسي تدرج ضمنه حقول معجمية دالة على المشاعر، والتعبير عن العواطف، والاحكام، والتقديرات، والسلوكات والتصرفات، وكذلك تقييم الأشياء وغيرها.

- **التدرج**: يضم هذا النظام فعل التثمينات التي تمثل خاصية عامة للتدرج، ويسمح بتفاوت شدة الخصائص المتعلقة بدرجة الإيجابية والسلبية المرتبطة بالحكم، هذه الخاصية تتوافق مع الفئات الثلاث للموقف المتمثلة في **"العاطفة والحكم والتقدير"**، كما يتوافق مع نظام الارتباط.

- **الارتباط**: وهو يحيل على دلالات المتحدثين الذين يعترفون أو ينكرون اختلاف وجهات النظر حسب مواقفهم الخاصة وانطلاقا من أماكن وجودهم. وقد اعتمد أحد الباحثين النظرية لتوضيح هذا الفرق على المفاهيم التي وضعها ميخايل بختين والمتمثلة في مونوكلوسيك (Monoglossic) وإيتيروكلوسيك (Heteroglossic)؛ إيتيروكلوسيك ويعمل على وصف وجود الخطاب بين ألسنة مختلفة ضمن سنن لساني واحد، وهذا الخطاب قد يشمل خطاب الشخصيات، خطاب الراوي، وخطاب الكاتب، في حين أن مونوكلوسيك يهتم بوجود الخطاب ضمن سنن واحد[34].

يضم نظام الموقف مجموعة من الحقول المعجمية، والمطلع على المبحث الذي خصصه الباحثان مارتن وويت لنظام الموقف في كتابهما مارتن وويت سيجد أنهما وضعا تصنيفا أوليا لهذه الحقول المعجمية (انظر الخطاطة 2)، وبعد قيامهما بتطبيقات على بعض النصوص تبين لهما أن بعض الحقول الفرعية يمكن تقسيمها إلى حقول فرعية أخرى (انظر الخطاطة 3)، إلا أن معظم الأبحاث الحالية تعتمد في تصنيف كلمات معجمها على التصنيفات الموجودة في الخطاطة 2، وتصنيف الكلمات لا يقف عند وضع الكلمة ضمن حقلها المعجمي فقط بل هناك معايير أخرى تتمثل في: تحديد قطبيتها ودرجة قوتها وتوجيهها.

(34) A. Amar و A. Lachkar, "Vers la Construction d'un Lexique d'Appraisal en Langue Arabe pour l'Analyse des Sentiments dans les Réseaux Sociaux", *Inst. D'Etudes Rech. Pour L'Arabisation Rabat,* نوفمبر 2015



الخطاطة 2: الحقول المعجمية لنظام الموقف المعتمدة في الأبحاث الحالية

7. خطوات بناء المعجم:

لتدارك ما تضمنته المعاجم المتخصصة بتحليل الآراء والمشاعر السالفة الذكر المترجمة إلى اللغة العربية من نقص على مستوى أقسام الكلام العربي – الذي يمكن توظيفه في التعبير عن الآراء والمشاعر وعدم الاكتفاء بالصفة –، جاءت فكرتنا في العمل على وضع معجم عربي يضم أقسام الكلام العربي تمثل بالأساس في مجموعة المشتقات انطلاقا من تعاريفها التي تمت الإشارة إليها في الكتب النحوية والبلاغية[35]،[36] والتي يمكن استعمالها في التعبير عن الآراء والمشاعر، وفيما يلي سنورد أهم الخطوات التي تم اتباعها في عملية بناء معجمنا والمتمثلة في ثلاث خطوات:

(35) M. A. K. Halliday وC. M. I. M. Matthiessen, *An introduction to functional grammar*, 3rd ed. London : New York: Arnold ; Distributed in the United States of America by Oxford University Press, 2004.

(36) ل. النجار، دور البنية الصرفية في وصف الظاهرة النحوية وتقعيدها، الأولى. الأردن، دار البشير، 1993.

الخطوة الأولى: الدراسة القبلية:

تتمثل الخطوة الأولى في دراسة المستوى الصرفي باعتباره المستوى اللساني الذي يهتم بتكوين ودراسة الكلمة العربية، فوجدنا أن الصرفيين العرب القدماء قد عرجوا على المباحث الدلالية التي تساعد في توجيه تشكيل الكلام العربي، بما فيها الأبنية الثلاثية المجردة والمزيدة، ومن بين الكتب المعتمدة في الدراسة الدلالية للأفعال نذكر:

- "**الكتاب**" لسيبويه أبي بشر عمرو بن عثمان بن قنبر.
- "**دروس في التصريف**" لمحمد محيي الدين عبد الحميد.
- "**أبنية الفعل دلالاتها وعلاقاتها**" لأبي أوس إبراهيم الشمسان.

لقد بينت هذه الكتب وغيرها أن كل وزن من الأوزان الثلاثية له دلالات معينة، يمكن توضيحها على الشكل الآتي: فبنية فعُل "بضم العين ولا يكون إلا لازما، ولا يجيء إلا في أفعال الغرائز والطبائع، نحو: أدُب، وأرُب، وجنُب، ..."[37]، وكذلك ما أشار إليه أبو أوس إبراهيم الشمسان حيث قال "أما ما كان حسنا أو قبحا فإنه مما يبنى فعله على فعُل يفعُل وذلك قولك: قبُح يقبُح، وسُم يوسُم، جمُل يجْمُل"[38]، أما بنية فعِل "بكسر العين، يجيء لازما ومتعديا، إلا أن لزومه أكثر من تعديه، ولذا غلب مجيء الأفعال على النعوت الملازمة وكبر الأعضاء ... ذرِب لسانه ... عرِج، وعوِج ..."، وبنية فعَل «بفتح العين، من أخف الأبنية، ولذا وضعوه للنعوت اللازمة، والأعراض، والأمراض، والألوان، ... نحو: بكَى، ... ستَر»[39].

(37) م. ع. ع. ا. الصيمري، "أبنية المشتقات في نهج البلاغة دراسة دلالية"، ماستر، كلية الآداب، جامعة البصرة، العراق، 2002.

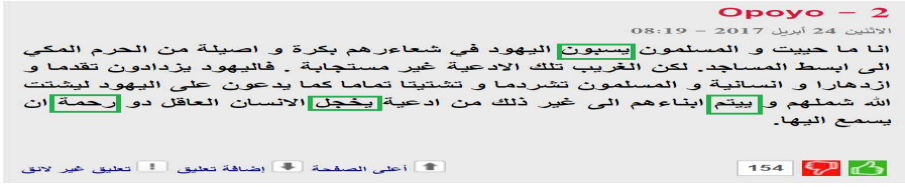
(38) ع. ا. محيي الدين، دروس التصريف. بيروت: المكتبة العصرية، 1995.

(39) أ. أ. إ. الشمسان، أبنية الفعل دلالاتها وعلاقاتها. المملكة العربية السعودية، دار المدني للطباعة والنشر والتوزيع، 1987.

ومن جهة أخرى، وبعد الاطلاع على مدونة التعاليق التي تم إنشاؤها من طرف معهد الدراسات والأبحاث للتعريب بالرباط والتي تضم أزيد من 100000 تعليق، تبين لنا أن الآراء والمشاعر يمكن التعبير عنها بالأسماء كما يمكن التعبير عنها كذلك بالأفعال؛ فالأسماء تتمثل في "المصادر، اسم المفعول والصفة المشبهة واسم الفاعل، وصيغة التفضيل"، ويمكن الإشارة إلى أمثلة من هذه التعاليق في الصور1 و2 و3.

الصورة 1: تعليق مأخوذ من موقع goodreads

الصورة 2: تعليق مأخوذ من موقع d'Aljazira



الصورة 3: تعليق مأخوذ من موقع Hespress

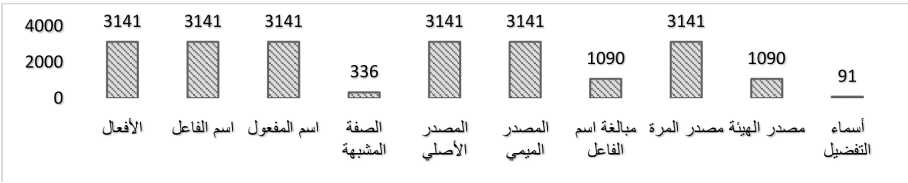
الخطوة الثانية: مرحلة الجمع:

قمنا بجمع مجموعة من الحروف التي من شأنها أن تغير قطبية الرأي شملت هذه الحروف: حروف النفي، وحروف النهي، وحروف الندبة... وضمن قاعدة

المعطيات التي يتوفر عليها معهد الدراسات والأبحاث للتعريب بالرباط والتي تضم 22856 فعل عربي، قمنا باختيار مجموعة من الأفعال التي يمكن اعتمادها في التعبير عن الآراء والمشاعر بالاعتماد على المعاجم التالية:

1. لسان العرب لابن منظور.
2. المعجم الوسيط من إخراج إبراهيم مصطفى وحامد عبد القادر وأحمد حسن الزيان ومحمد علي النجار.
3. معجم العين للخليل بن أحمد الفراهيدي.

مكنتنا هذه العملية من الحصول على 3141 فعلا مجردا ومزيدا، أما مشتقات هذه الأفعال فقد تم الحصول عليها باعتماد برنامج اشتقاقي أنجز بمعهد الدراسات والأبحاث للتعريب، والرسم البياني يوضح الأرقام المحصل عليها بعد عملية الاشتقاق بشكل مفصل:



يوضح الرسم البياني مجموع الكلمات المشتقة حسب فئتها النحوية؛ نتج عن هذه العملية تساوي كل من: "اسم الفاعل، واسم المفعول، والمصدر الأصلي، والمصدر الميمي، ومصدر المرة، وبلغ كل منها 3141، في حين تساوى كل من "مبالغة اسم الفاعل ومصدر الهيئة"، وبلغ عدد كل منهما 1090، وتبقى الصفة المشبهة أقل المشتقات لأسباب نحوية. أما فيما يخص أسماء التفضيل فقد اكتفينا بجمعها من كتاب "أفعل من كذا" أبو علي القالي وبلغ عددها 91[40].

الخطوة الثالثة: مرحلة التصنيف

بعد الوقوف في بداية البحث عند التطور الذي عرفته مختلف الدراسات المعجمية وما نتج عنها من إصدار لمجموعة من المعاجم المتخصصة التي

(40) إ. ب. ا. ب. ع. ا. ب. ع. القالي، أفعل من كذا، الأولى. دمشق، دار سعد الدين، 1431.

ساعدت في توجيه الكاتب الوجهة السليمة في اختيار الكلمات الملائمة تبعا للسياق الذي عبر من خلاله عن آرائه أو مشاعره، ارتأينا الاعتماد على بعضها، خاصة معجم "نجعة الرائد وشرعة الوارد في المترادف والمتوارد" لإبراهيم اليازجي اللبناني، (الخطاطة 4) الذي أخذنا منه سبعة حقول معجمية، وهي: "التباس الأمر/ اتضاحه، اتفاق الرأي، سهولة الأمر/ صعوبة الأمر، الشجاعة والجبن، قوة البنية/ ضعف البنية، حسن المنظر/ قبحه، والخلق"،[41] كما اعتمدنا معجم و"المكنز الكبير" لأحمد مختار عمر الذي اقتصرنا فيه فقط على حقلين اثنين هما: "الإيمان/ الكفر، والإرهاب/ التهدئة"[42].

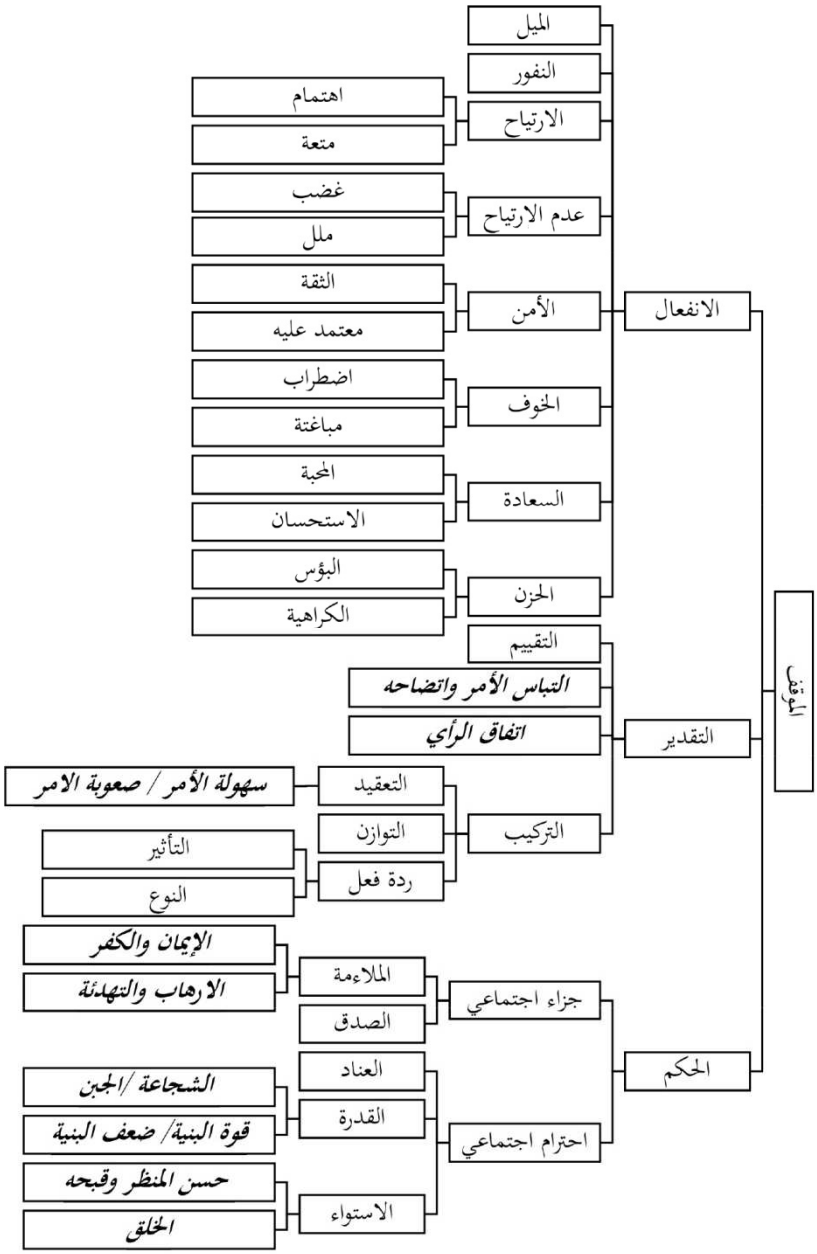
وفيما يلي التصنيف النهائي لكلمات المعجم وفق الحقول المعجمية المذكورة سلفا، مع توسيم كل كلمة بالسمات الموالية:

- تحديد جذع الكلمة.
- تحديد الخاصية الصرفية لكل كلمة حسب المشتقات التي تم اختيارها، وهي: اسم الفاعل، اسم المفعول، الصفة المشبهة، مبالغة اسم الفاعل، المصدر الأصلي، المصدر الميمي، مصدر المرة، ومصدر الهيئة.
- تحديد توجيه الكلمة لمعرفة ما إذا كانت موجبة أو سالبة أو محايدة؛ ففي عملية التصنيف نجد في مشتقات الفعل الواحد منها الموجبة، والسالبة، والمحايدة.
- تحديد درجة قوة الكلمة؛ حيث نجد الكلمات قوية، ومتوسطة، وضعيفة، ويتم هذا التوسيم بالاعتماد على معاني الكلمات في المعاجم المعتمدة.
- تحديد قطبية الكلمة، ونشير هنا إلى أن كل كلمات المعجم لها نفس القطبية **"غير محددة"**(unmarked)، حيث يتم تغييرها أثناء عملية التحليل من **"غير محددة"** (unmarked) إلى "محددة" (marked) إذا كان التعبير يشتمل على أحد الأدوات التي تغير معنى الكلمة من الموجب إلى السالب أو العكس،

(41) إ. اليازجي اللبناني، نجعة الرائد وشرعة الوارد في المترادف والمتوارد، بيروت، مكتبة لبنان، 1985.

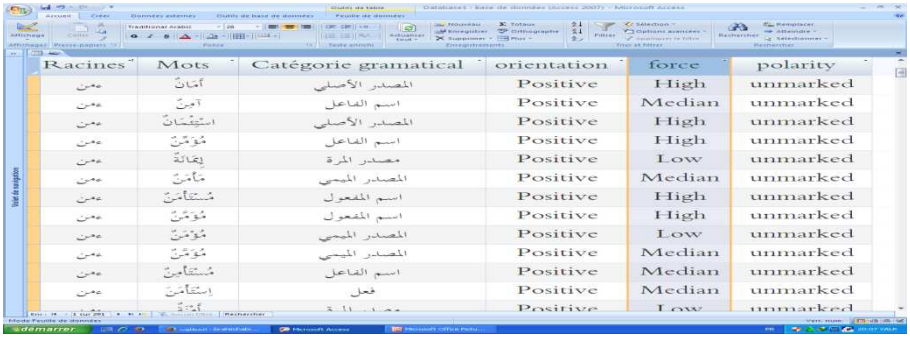
(42) أ. مختار عمر، المكنز الكبير معجم شامل للمجالات والمترادفات والمتضادات، الأولى، المملكة العربية السعودية، مؤسسة مارينا لخدمات الطباعة.

وللمزيد من التوضيح نقترح ثلاثة أمثلة من معجمنا تضم تصنيفا لبعض الكلمات وفق التوسيمات السابقة الذكر.



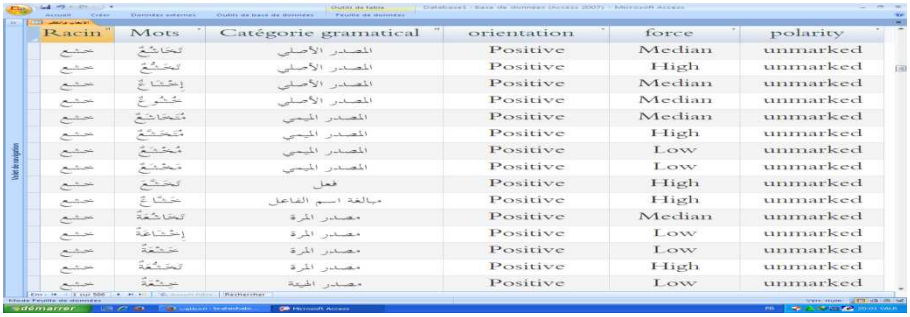
الخطاطة 4: الحقول المعجمية لنظام الموقف بعد إضافة بعض الحقول

- **الانفعال: حقل الثقة**



Racines	Mots	Catégorie gramatical	orientation	force	polarity
مءن	أَمَانٌ	المصدر الأصلي	Positive	High	unmarked
مءن	آمِنٌ	اسم الفاعل	Positive	Median	unmarked
مءن	اسْتِئْمَانٌ	المصدر الأصلي	Positive	High	unmarked
مءن	مُؤَمِّنٌ	اسم الفاعل	Positive	High	unmarked
مءن	مِئْمَانَةٌ	مصدر المرة	Positive	Low	unmarked
مءن	مَأْمَنٌ	المصدر الميمي	Positive	Median	unmarked
مءن	مُسْتَأْمَنٌ	اسم المفعول	Positive	High	unmarked
مءن	مُؤَمَّنٌ	اسم المفعول	Positive	High	unmarked
مءن	مُؤْمَنٌ	المصدر الميمي	Positive	Low	unmarked
مءن	مُؤَمَّنٌ	المصدر الميمي	Positive	Median	unmarked
مءن	مُسْتَأْمِنٌ	اسم الفاعل	Positive	Median	unmarked
مءن	اسْتَأْمَنَ	فعل	Positive	Median	unmarked
مءن	أَمَنَةٌ	مصدر المرة	Positive	Low	unmarked

- **الحكم: حقل الايمان والكفر**



Racin	Mots	Catégorie gramatical	orientation	force	polarity
خشع	لَخَاشِعٌ	المصدر الأصلي	Positive	Median	unmarked
خشع	لَخَشْعٌ	المصدر الأصلي	Positive	High	unmarked
خشع	إِخْشَاعٌ	المصدر الأصلي	Positive	Median	unmarked
خشع	خُشُوعٌ	المصدر الأصلي	Positive	Median	unmarked
خشع	مُتَخَاشِعٌ	المصدر الميمي	Positive	Median	unmarked
خشع	مُتَخَشَّعٌ	المصدر الميمي	Positive	High	unmarked
خشع	مُخْشَعٌ	المصدر الميمي	Positive	Low	unmarked
خشع	مَخْشَعٌ	المصدر الميمي	Positive	Low	unmarked
خشع	لَخَشَّعَ	فعل	Positive	High	unmarked
خشع	خَشَّاعٌ	مبالغة اسم الفاعل	Positive	High	unmarked
خشع	لَخَاشَعَةٌ	مصدر المرة	Positive	Median	unmarked
خشع	إِخْشَاعَةٌ	مصدر المرة	Positive	Low	unmarked
خشع	خَشْعَةٌ	مصدر المرة	Positive	Low	unmarked
خشع	لَخَشَّعَةٌ	مصدر المرة	Positive	High	unmarked
خشع	خِشْعَةٌ	مصدر الهيئة	Positive	Low	unmarked

- **التقدير: حقل اتفاق الرأي**

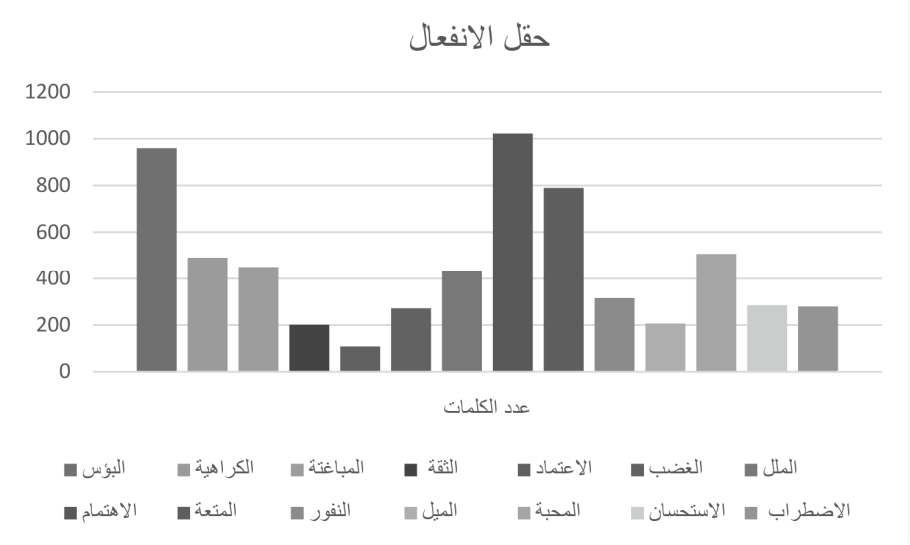
Racines	Mots	Catégorie gramatical	orientation	force	polarity
عيد	تَأْيِيدَةٌ	مصدر المرة	Positive	low	unmarked
عيد	تَأْيِيدٌ	المصدر الأصلي	Positive	median	unmarked
عيد	تَأْيِيدٌ	المصدر الأصلي	Positive	median	unmarked
عيد	مُؤَيِّدٌ	المصدر الميمي	Positive	high	unmarked
عيد	مُتَأَيَّدٌ	اسم المفعول	Positive	median	unmarked
عيد	مُؤَيِّدٌ	اسم الفاعل	Positive	low	unmarked
عيد	مُتَأَيِّدٌ	المصدر الميمي	Positive	median	unmarked
عيد	أَيَّدَ	فعل	Positive	high	unmarked
عيد	مُؤَيَّدٌ	اسم المفعول	Positive	high	unmarked
عيد	تَأْيِيدَةٌ	مصدر المرة	Positive	low	unmarked
عيد	مُتَأَيِّدٌ	اسم الفاعل	Positive	median	unmarked
بخس	مُبَخَّسٌ	المصدر الميمي	Negative	median	unmarked
بخس	مُبَخَّسٌ	اسم المفعول	Negative	median	unmarked

8. النتائج والأبحاث المستقبلية:

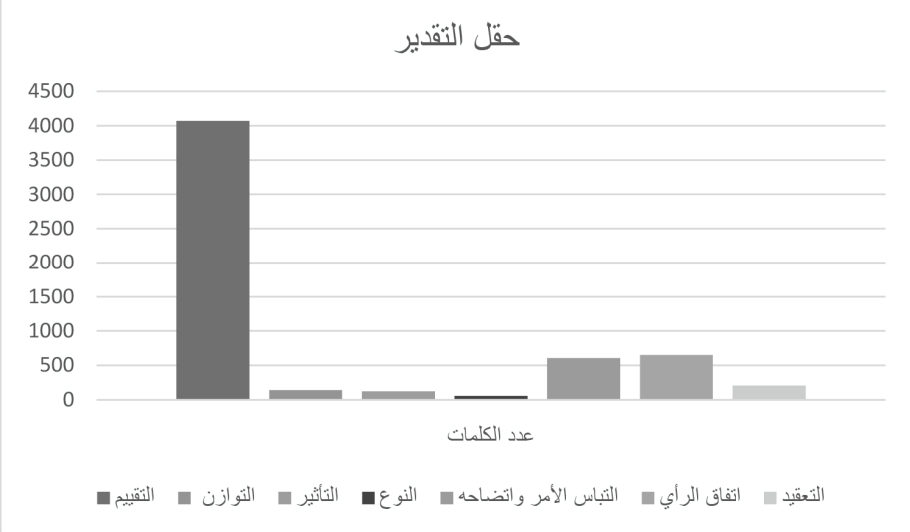
أ. النتائج:

عملنا في هذا البحث على تقديم الخطوات المتبعة في بناء ووسم معجم عربي اعتمادا على نظرية الأبريسال؛ إذ حاولنا من خلاله تدارك النقص الحاصل في المعاجم المتوفرة التي تُرجمت إلى اللغة العربية باعتماد الترجمة الآلية، كمقاربة لبناء معاجم موجهة لتحليل الآراء والمشاعر باللغة العربية مع إغفال خصوصية كل لغة، أضف إلى ذلك أن تلك المعاجم اقتصرت على الصفات فقط في حين أنها أغفلت باقي أقسام الكلام العربي، من هذا المنطلق ارتكزت فكرتنا في بناء معجم عربي موجه لتحليل الآراء والمشاعر على اعتمادنا ثلاث خطوات أساسية؛ تمثلت الأولى في الدراسة القبلية التي شملت الاطلاع على المباحث الصرفية وما لها من دور في بناء المعجم العربي، أما الثانية فتجلت في جرد 3141 من الأفعال الثلاثية المجردة والمزيدة التي يمكن استعمالها في التعبير عن الأحاسيس والآراء والمشاعر من قاعدة معطيات يتوفر عليها معهد الدراسات والأبحاث للتعريب بالرباط ضمت أكثر من 22000 فعل، وانتقلنا بعد ذلك لإغناء المعجم بالاعتماد على عملية الاشتقاق التي وظفنا فيها برنامجا اشتقاقيا مكننا من الحصول على 21453 كلمة، ثم انتقلنا إلى تصنيف هذه الكلمات وفق الحقول المعجمية التي جاءت بها نظرية الأبريسال، والاعتماد كذلك على الموروث العربي القديم منه والحديث لإضافة حقول معجمية أخرى، وسنورد رسوم بيانية توضح نسب الحقول المعجمية الثلاثة الرئيسة:

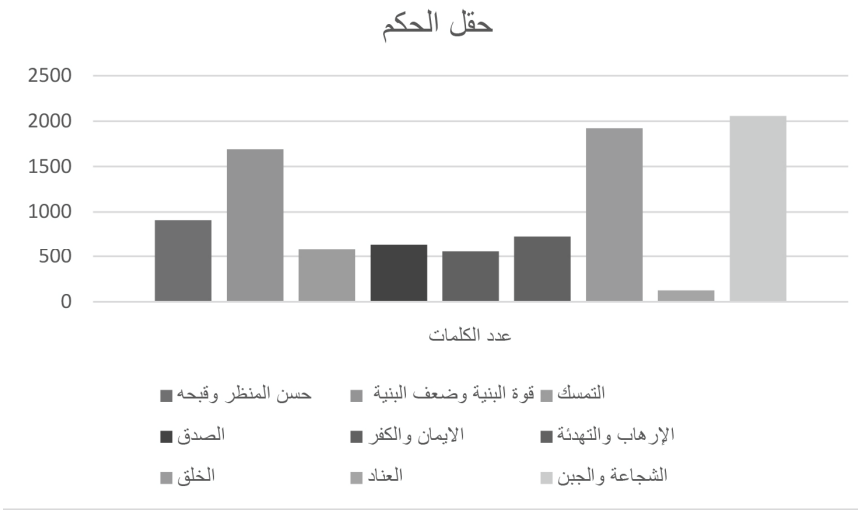
- **حقل الانفعال:**



- **حقل التقدير:**



o حقل الحكم:



ب. **الأبحاث المستقبلية**

يقتضي البحث في مجال تحليل الآراء والمشاعر التوفر إلى جانب المعجم على مدونة تضم تعاليق يتم جمعها من مواقع مختلفة لهذا الغرض، ونحن نعمل على وسم المدونة التي يتوفر عليها معهد الدراسات والأبحاث للتعريب بالرباط وتضم تعاليق مكتوبة باللغة العربية تم تحميلها من مواقع مختلفة تضم أزيد من 100000 تعليق، كما نعمل بموازاة ذلك على تنمية المعجم الذي بيناه بالأفعال الرباعية المجردة منها والمزيدة باتباع نفس الخطوات.

9. خاتمة:

تعتبر الدراسات المعجمية من أهم الدراسات اللغوية التي حظيت باهتمام مجموعة من الباحثين الرواد من جوانب متعددة ساعدت على ظهور ما سمي "بالمعاجم المتخصصة"، التي تهدف إلى إفهام الباحث وتعريفه بكلمات مجال بحثه، كما أن اللسانيات الحاسوبية ساهمت بشكل كبير في تطوير هذا النوع من المعاجم، إضافة إلى أن التطورات المتعلقة بمجال المعلوماتية التقانية (والتكنولوجية) والويب

ساهمت في تنمية وتطوير المواقع الإلكترونية، الشيء الذي مهد الطريق لظهور مجال تحليل الآراء والمشاعر، الأمر الذي فرض توفير معاجم آلية تتولى تسهيل عملية تحليل الآراء ومعالجتها. ونظرا لأن المعاجم المتوفرة في هذا المجال معاجم إنجليزية والتي تم ترجمتها آليا، وتطبيقها في العديد من اللغات من بينها اللغة العربية دون مراعاة لخصوصية كل لغة على حدة.

كل هذا دفعنا إلى التفكير في بناء معجم عربي متخصص في تحليل الآراء والمشاعر انطلاقا من نظرية الأبريسال اللسانية في تصنيف كلمات المعجم مع العودة إلى الموروث المعجمي والصرفي للغة العربية من جهة، ومن جهة أخرى استغلال الأنظمة الحاسوبية في إغناء هذا المعجم بالمشتقات.

وتجدر الإشارة إلى أن الفريق الذي يأخذ على عاتقه مهمة المعالجة الآلية للغة العربية التابع لمعهد الدراسات والأبحاث للتعريب بالرباط يشتغل حاليا على إنشاء برنامج معلوماتي لتحليل الآراء والمشاعر المكتوبة باللغة العربية، كما نشير كذلك إلى أننا بصدد إغناء معجمنا بالجذوع الرباعية واشتقاقاتها.